

# Entropic information–disturbance tradeoff

Lorenzo Maccone

*QUIT - Quantum Information Theory Group, Dipartimento di Fisica  
“A. Volta” Università di Pavia, via A. Bassi 6, I-27100 Pavia, Italy.*

We show the flaws found in the customary fidelity-based definitions of disturbance in quantum measurements and evolutions. We introduce the “entropic disturbance”  $D$  and show that it adequately measures the degree of disturbance, intended essentially as an irreversible change in the state of the system. We also find that it complies with an information–disturbance tradeoff, namely the mutual information between the eigenvalues of the initial state and the measurement results is less than or equal to  $D$ .

PACS numbers: 03.67.-a, 03.65.Yz, 03.65.Ta, 42.50.Lc

The correct interpretation of the Heisenberg uncertainty relations [1, 2] connects the uncertainty (or spread) in the measurement results of one observable to the spread of another observable in the initial state of the system [3]. The true spirit of Heisenberg’s intuition [1, 4], however, was that any measurement (i.e. extraction of classical information from a quantum system) necessarily entails some kind of disturbance to the measured system. This intuition has yet to be captured in a truly general information–disturbance tradeoff relation. Nonetheless, several of these relations have been put forth and cover many conceivable situations (e.g. see [4, 5, 6, 7]). The main problem in deriving a general version of such a tradeoff lies in the identification of an appropriate definition for the disturbance.

In this paper we explore the consequences of defining the disturbance as an irreversible change in the state of the system. We proceed axiomatically, by enumerating the abstract properties that a disturbance measure should possess, and we show that the customarily used fidelity-based disturbance measures do not satisfy them. We introduce one that does, namely the “entropic disturbance”  $D$ . A simple, general information–disturbance tradeoff is then derived, namely  $I \leq D$  ( $I$  being the mutual information between the eigenvalues of the initial state and the measurement results). Its interpretation is straightforward: Every time an amount  $I$  of information is obtained from the measurement apparatus, a disturbance  $D$  at least as big is introduced on the system, but, obviously, the system can be disturbed by a process that returns little information. We also derive the equality conditions  $I = D$  showing that the  $I \leq D$  bound is tight, and proving (as is to be expected [8]) that von Neumann-type measurements are among the least disturbing amid all measurements that retrieve the same information. Finally, we give a state-independent tradeoff by averaging over all possible input states. This provides a measure of the global disturbance introduced by the apparatus or by a generic evolution.

## I. MEASUREMENT

A measurement is by definition an operation that acts on a system and returns some classical information, i.e. a label “ $l_k$ ” identified by an index  $k$ . This operation typically, but not always, changes the state of the measured system (*wave-function collapse* mechanism). When measuring classical objects, the state change is only due to the change in the information we have on the system: The uncertainty in the state is usually reduced. When measuring quantum objects, the state change can have a dynamical nature that perturbs the system. The rigorous description of these mechanisms results from the Kraus decomposition of the measurement apparatus [6, 9], which identifies the completely positive (CP) map of its evolution. When the system is initially in a state  $\rho$ , the  $k$ th measurement outcome occurs with a probability  $p_k = \text{Tr}[\Pi_k \rho]$ , where  $\{\Pi_k\}$  is the apparatus POVM (Positive Operator-Valued Measure), a set of positive operators normalized so that  $\sum_k \Pi_k = \mathbb{1}$ . After the outcome  $l_k$  is obtained, the state is changed to

$$\rho'(k) = \sum_{i \in I_k} K_i \rho K_i^\dagger / p_k, \quad (1)$$

where  $I_k$  is a set of indices  $i$  and  $K_i$  are the apparatus Kraus operators. The POVM in terms of these is given by  $\Pi_k = \sum_{i \in I_k} K_i^\dagger K_i$ . On the basis of the above description of the measurement, we can define the following information and disturbance measures.

*Information:* The outcome  $l_k$  provides the experimenter with some information  $I$ . We are, obviously, interested in the case in which  $l_k$  provides some information on the measured system, and it is not just independently generated by the apparatus. Thus, a good measure for  $I$  is the mutual information between the measurement results and some property of the system state. A significant (basis-independent) property is the spectrum of the state  $\rho$ , i.e. the probability distribution of its eigenvalues  $\lambda_j$ . Of course, the ensemble composed by the eigenvectors of  $\rho$  weighted by the corresponding eigenvalues is not the only one that originates from the state  $\rho$ . However, it is easy to see that it is the one that allows to recover the maximal accessible information (it saturates the Holevo

bound) [10]. It is thus the appropriate choice since we ultimately want to bound the information from above. We then use

$$I \equiv I(\lambda_j, p_k), \quad (2)$$

the mutual information between the label  $j$  identifying the eigenvalue of  $\rho$  and the label  $k$  identifying the measurement result:  $I$  is the number of bits that the experimenter gains from the result  $l_k$  on which eigenvector of  $\rho$  the system was in (before the measurement). It is maximal when all eigenvalues are equal (the experimenter has no prior info on the state), and it is null when the state is pure (the experimenter already has total knowledge of the state, and cannot gain any more information on it). [Note that  $\rho$  here refers to the state from the experimenter's point of view: It reflects his prior knowledge of the system state, i.e. that each eigenvector has a prior probability  $\lambda_j$ . The 'true' state (i.e. from the point of view of who is preparing the system) will be in general purer. The knowledge that can be acquired by the experimenter is upper bounded by the difference in entropy between these two representations of the state.] If the Hilbert space of the system has finite dimension  $d$ , we can normalize  $I$  by dividing it with its maximum value  $\log_2 d$ , so to have  $0 \leq I \leq 1$ .

*Disturbance:* A disturbance is an irreversible change in the state of the system, caused by a CP-map evolution (such as the dynamical disturbance caused by quantum correlations that leak out to the environment and are lost). Thus, any quantity  $D$  that measures disturbance should satisfy the following requirements, inspired by Ref. [4]:

- i)  $D$  should be a function only of the input state  $\rho$  and of the apparatus, identified through its Kraus operators  $\{K_i\}$ , i.e.  $D = D(\rho, \{K_i\})$ .
- ii)  $D$  should be null if and only if the transformation  $\{K_i\}$  is invertible on  $\rho$  [22]. In this case the state change can be undone, and such transformation is not disturbing the system.
- iii) Once the state has been disturbed, it should not be possible to decrease  $D$  with any successive transformation. This means that  $D$  should be monotonically non-decreasing for successive applications of CP-maps [8] (i.e. it should satisfy a sort of pipeline inequality). This requirement captures the notion that a disturbance should be irreversible, and is connected with the concept of "cleanness" [11].
- iv)  $D$  should be continuous: Maps and input states which do not differ too much should give similar values of  $D$ .

The above requirements, which define the disturbance axiomatically, have nothing to do with the information the measurement provides. As such, there is no obvious *a priori* reason why an information-disturbance tradeoff should hold.

Definitions of disturbance are customarily based on the fidelity or the Bures distance [12] between input and output states. Even though valid information-disturbance relations can be found [5], these definitions do not seem to appropriately gauge the disturbance, intended as an irreversible evolution. In fact, even though a unitary transformation is perfectly reversible, it can rotate a state to an orthogonal configuration, generating the maximum possible fidelity-based disturbance. These quantities do not satisfy the requirements *ii*) and *iii*). Analogous considerations apply also if we use the entanglement fidelity [13] in place of the fidelity [23].

A definition of disturbance  $D$  that satisfies all the above requirements can be found by recalling that a CP-map  $\mathcal{Q}$  is invertible if and only if [14] the map's coherent information  $I_c(\rho, \mathcal{Q})$  is equal to the von Neumann entropy  $S(\rho) = -\text{Tr}[\rho \log_2 \rho]$  of the input state  $\rho$ . The coherent information [14, 15] is defined as  $I_c \equiv S(\mathcal{Q}(\rho)) - S((\mathcal{Q} \otimes \mathbb{1})(|\Psi\rangle\langle\Psi|))$ , where  $|\Psi\rangle$  is a purification of  $\rho$  and the map  $\mathcal{Q} \otimes \mathbb{1}$  acts with  $\mathcal{Q}$  on the system space and with the identity  $\mathbb{1}$  on the purification space. The quantity  $I_c$  is non-increasing for application of CP-maps (data-processing inequality) [14]. Namely, for any two maps  $\mathcal{Q}$  and  $\mathcal{Q}'$ , we have  $I_c(\rho, \mathcal{Q}) \geq I_c(\rho, \mathcal{Q}' \circ \mathcal{Q})$ , where  $\circ$  denotes composition of maps. Thus, a disturbance measure that satisfies requirements *ii*)-*iii*) must be a function  $f$  of  $S(\rho) - I_c$ , with  $f$  non-decreasing and null when its argument is:  $f(0) = 0$ . We then define

$$\begin{aligned} D &\equiv S(\rho) - I_c \\ &= S(\rho) - S(\mathcal{Q}(\rho)) + S((\mathcal{Q} \otimes \mathbb{1})(|\Psi\rangle\langle\Psi|)), \end{aligned} \quad (3)$$

which, in addition to *ii*)-*iii*), also satisfies requirements *i*) and *iv*) since it is continuous (see the Appendix). Analogously to  $I$ , also  $D$  can be normalized in  $d$ -dimensional Hilbert spaces by dividing it by  $\log_2 d$ , so that  $0 \leq D \leq 2$ .

From the postulates of quantum mechanics it follows that the system state  $\rho$  describes the information the experimenter possesses on the system. Hence, there are two mechanisms that lead to a state change: the system dynamics and the acquisition of new information. Heisenberg, in his uncertainty principle, was considering the former mechanism only. To exemplify the latter, suppose I acquire a qubit in an unknown state. Initially, I will assign to it the state  $\mathbb{1}/2$ , but as soon as the preparer tells me that the qubit was in the state  $\rho_p$ , from my point of view it undergoes a state change (even though I may have not interacted with it) described by the map  $\rho_p = \mathcal{C}[\mathbb{1}/2]$ . We can call this a "purely informational" state change. Since both the dynamical and the informational state changes are described by CP-maps, they both fall in the general framework described above. [In this sense it may be interpreted as a generalization of Heisenberg's intuition]. Is it possible to weight the contribution of these two mechanisms in each measurement apparatus? Yes: Since the set of CP-maps is a convex set, the apparatus CP-map  $\mathcal{Q}$  (identified by the Kraus operators  $\{K_i\}$ ) can always be written as a convex combination of the purely informational map  $\mathcal{C}$  and of a "dynamical"

map  $\mathcal{T}$  as

$$\mathcal{Q} = \xi\mathcal{C} + (1 - \xi)\mathcal{T} \text{ with } \xi \in [0, 1]. \quad (4)$$

The Kraus operators of the map  $\mathcal{C}$  are  $A_{jk} \equiv \sqrt{\mu_j}|v_j\rangle\langle v_k|$ , where  $\mu_j$  and  $|v_j\rangle$  are the eigenvalues and eigenvectors of the ‘true’ (i.e. from the point of view of the preparer) state  $\rho_p$ : The action of  $\mathcal{C}$  must not change the true state,  $\mathcal{C}[\rho_p] = \rho_p$ . [Note that in the case of a degenerate state  $\rho_p$ , the map  $\mathcal{C}$  might not be univocally defined]. The POVM of a purely informational measurement is then  $\Pi_k = \sum_j A_{jk}^\dagger A_{jk} = |v_k\rangle\langle v_k|$  (the projectors on the eigenspaces of  $\rho_p$ ). In this respect, the *truly quantum* contribution to the disturbance in a measurement is related to the map  $\mathcal{T}$ : it is present in those measurement apparatuses with  $\xi < 1$ .

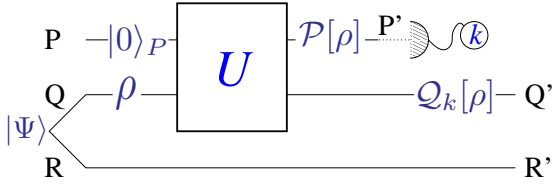


FIG. 1: Measurement apparatus described through the indirect measurement model. Any apparatus can be described [6] by coupling the system to be measured (in the space Q) with a probe P through a unitary operator  $U$ . The probe is then projectively measured at the output  $P'$ . Notation employed below:  $|0\rangle_P$  = initial state of the apparatus in P;  $\rho$  = initial state of the system in Q;  $\mathcal{P}[\rho]$  = state of the probe in  $P'$ , before the final projective measurement;  $\mathcal{Q}_k[\rho]$  = state of the system in  $Q'$ , after the interaction with the apparatus and after the probe measurement with result  $l_k$  (i.e. after the wave-function collapse). The reference R is introduced for purification purposes: It is defined so that the joint initial state  $|\Psi\rangle$  of system and reference, QR, is pure.

## II. INFORMATION-DISTURBANCE TRADEOFF

We now prove the information-disturbance tradeoff

$$I(\lambda_j, p_k) \leq D(\rho, \{K_i\}). \quad (5)$$

Any measurement apparatus can be decomposed into a unitary evolution  $U$  (the Stinespring dilation of the apparatus) followed by a von Neumann projective measurement on a probe P, the so-called *indirect measurement model* [6], see Fig. 1. The unitary  $U$  couples Q with the probe P in the apparatus yielding  $Q'$  and  $P'$ . The joint evolution of probe and system PQ in the apparatus can be seen as composed by two complementary quantum channels: A channel  $Q \rightarrow P'$  that describes the transfer of information from the system to the state of the probe in  $P'$  which is then measured yielding the measurement result  $l_k$ , and a channel

$Q \rightarrow Q'$  that evolves the system before the measurement into the system after the measurement conditioned on its result  $l_k$ . These two channels are respectively described by the CP-maps  $\mathcal{P}[\rho] \equiv \text{Tr}_Q[U(\rho \otimes |0\rangle_P\langle 0|)U^\dagger]$  and  $\mathcal{Q}_k[\rho] \equiv \text{Tr}_P[U(\rho \otimes |0\rangle_P\langle 0|)U^\dagger(\mathbb{1}_Q \otimes |k\rangle_P\langle k|)]/p_k$ , where  $|0\rangle_P$  is the initial pure state of the probe,  $|k\rangle_P$  is the basis representing the projective measurement on the probe, and  $p_k$  is the probability of the  $k$ th result. Since the final state of the system in  $Q'$  is given by  $\rho'(k) = \mathcal{Q}_k[\rho]$  with probability  $p_k$ , it can be written as  $\sum_k p_k \rho'(k) = \text{Tr}_P[U(\rho \otimes |0\rangle_P\langle 0|)U^\dagger] \equiv \mathcal{Q}[\rho]$ , where the Kraus operators of the map  $\mathcal{Q}$  can be immediately obtained from the ones of the maps  $\mathcal{Q}_k$ . The map  $\mathcal{Q}$  describes the unitary coupling of system and probe in the apparatus and the successive trace on the probe space, which yields the unconditioned output state.

The system's initial state  $\rho$ , expanded on its eigenvectors  $|j\rangle$  is given by  $\rho = \sum_j \lambda_j |j\rangle\langle j|$ . We use the Holevo-Schumacher-Westmoreland theorem [16] with an alphabet composed by  $|j\rangle\langle j|$  with probability  $\lambda_j$  flowing through a channel described by  $\mathcal{P}$ . Such theorem implies that the mutual information  $I(\lambda_j, p_k)$  between the variable  $j$  and the measurement results  $k$  (whatever measurement strategy is employed) is upper bounded as

$$I(\lambda_j, p_k) \leq S(\mathcal{P}[\rho]) - \sum_j \lambda_j S(\mathcal{P}[|j\rangle\langle j|]). \quad (6)$$

The system in space Q can be purified by adding an auxiliary reference space R, so that the system in QR is initially in a pure state  $|\Psi\rangle$ . The entropy  $S(P') \equiv S(\mathcal{P}[\rho])$  of the probe just before the final von Neumann measurement is then equal to the entropy  $S(Q'R')$  of the joint state in  $Q'R'$  of the system and the reference after the interaction  $U$ . In fact, the initial state in PQR is pure and it is evolved into  $P'Q'R'$  by a unitary evolution. Thus,

$$S(\mathcal{P}[\rho]) = S(Q'R') = S_e(\rho, \mathcal{Q}), \quad (7)$$

where  $S_e(\rho, \mathcal{Q})$  is the exchange entropy [13, 15] of the map  $\mathcal{Q}$ . It is defined as the entropy of the joint  $Q'R'$  output state of system and purification-reference, i.e.  $S_e(\rho, \mathcal{Q}) \equiv S((\mathcal{Q} \otimes \mathbb{1}_R)[|\Psi\rangle\langle\Psi|])$ . Moreover, the entropy in  $Q'$  satisfies

$$\begin{aligned} S(Q') &\equiv S(\mathcal{Q}[\rho]) \\ &= S\left(\sum_j \lambda_j \mathcal{Q}[|j\rangle\langle j|]\right) \leq \sum_j \lambda_j S(\mathcal{Q}[|j\rangle\langle j|]) + H(\lambda_j), \end{aligned} \quad (8)$$

where  $H(\lambda_j) = S(\rho)$  is the Shannon entropy of the probability distribution  $\lambda_j$ , and where the inequality  $S(\sum_x p_x \rho_x) \leq \sum_x p_x S(\rho_x) + H(p_x)$  (valid for all probabilities  $p_x$  and states  $\rho_x$ ) has been used [12]. Notice that if the system in Q is initially in a pure state  $|j\rangle$ , the entropy of the output of the two channels  $\mathcal{P}$  and  $\mathcal{Q}$  coincides since the entropy of the joint system PQ is initially null. Hence,  $S(\mathcal{Q}[|j\rangle\langle j|]) = S(\mathcal{P}[|j\rangle\langle j|])$ , so that Eq. (8) implies

$$\sum_j \lambda_j S(\mathcal{P}[|j\rangle\langle j|]) \geq S(\mathcal{Q}[\rho]) - S(\rho). \quad (9)$$

Joining Eqs. (6), (7) and (9), we find  $I \leq S_e(\rho, \mathcal{Q}) - S(\mathcal{Q}[\rho]) + S(\rho) = S(\rho) - I_c(\rho, \mathcal{Q}) = D$ , thus proving Eq. (5). Notice that such proof works also in the case in which the input and output Hilbert spaces  $\mathcal{Q}$  and  $\mathcal{Q}'$  do not coincide, and in the case of infinite dimensional Hilbert spaces [17].

There is a simple, not-very-rigorous, intuition behind the preceding proof. The total quantum information of the initial state  $\rho$  can be quantified by  $N \simeq S(\rho)$  qubits. The unitary evolution  $U$  transfers  $n$  of them to the probe space  $\mathcal{P}$ , where the projective measurement can return a number of bits  $b \leq n$ , due to the Holevo bound [12]. The remaining  $N - n$  qubits constitute an upper bound to the quantum capacity to transfer the quantum information in the initial state through the channel  $\mathcal{Q} \rightarrow \mathcal{Q}'$  consisting of the measurement apparatus. The quantum capacity is measured by the coherent information [15, 18], so that  $I_c \leq N - n$ . Thus,

$$I \simeq b \leq n = N - (N - n) \leq N - I_c \simeq D. \quad (10)$$

We now deduce the equality conditions for the information–disturbance bound  $I \leq D$ , showing that it is achievable. The equality in the Holevo–Schumacher–Westmoreland relation of Eq. (6) is achieved if the alphabet states  $\mathcal{P}[|j\rangle\langle j|]$  commute [19]. Moreover, the equality in the relation  $S(\sum_x p_x \rho_x) \leq \sum_x p_x S(\rho_x) + H(p_x)$ , which was employed in Eq. (8), is achieved if and only if the states  $\rho_x$  have support on orthogonal subspaces [12] (which implies that they commute). Thus, we have equality  $I = D$  if and only if the channel  $\mathcal{P}$  maps different eigenvectors  $|j\rangle$  of the initial state  $\rho$  into orthogonal subspaces. A typical example is a projective measurement whose Kraus operators are projectors on the basis  $|j\rangle$ . It is a purely informational measurement, where the only uncertainty derives from classical probability. [Notice that, in a  $d$ -dimensional Hilbert space, the converse also holds for measurements with  $d$  outcomes: If  $I = D$  and the measurement POVM has  $d$  elements, then the measurement is a von Neumann-type projection, i.e. its Kraus operators are of the form  $A_j = \tilde{U}|a_j\rangle\langle a_j|$  where  $\tilde{U}$  is a fixed unitary and  $|a_j\rangle$  is a basis. The proof of this assertion follows immediately from the fact that the measurement must map a basis  $|j\rangle$  of  $d$  elements into  $d$  orthogonal subspaces, which, in a  $d$ -dimensional Hilbert space, must then be one-dimensional].

### III. STATE-INDEPENDENT TRADEOFF

The definitions we used for information  $I(\lambda_j, p_k)$  and disturbance  $D(\rho, \{K_i\})$  are explicitly dependent both on the input state  $\rho$  and on the apparatus. We can forgo the state dependence by averaging on all possible input states with equal weights (for symmetry reasons), i.e. by using a state  $\rho = \mathbb{1}/d$  in a  $d$ -dimensional Hilbert space. [In infinite-dimensional spaces, additional requirements are also necessary, such as using states with upper-bounded

energy.] Thus, we can define a state-independent information as  $\tilde{I}(p_k) \equiv I(\mathbb{1}/d, p_k)$  and a state-independent disturbance as  $\tilde{D}(\{K_i\}) \equiv D(\mathbb{1}/d, \{K_i\}) = \log_2 d - I_c(\mathbb{1}/d, \mathcal{Q})$ . The quantity  $\tilde{I}/\log_2 d$  measures the percentage of the maximum retrievable information that is achieved by the apparatus. The quantity  $\tilde{D}$  measures the disturbance the apparatus causes to a completely unknown state. Notice that  $\tilde{D} = 0$  if and only if the apparatus acts on the state with a unitary transformation (i.e. it yields no information on the state). In fact,  $S(\rho) = I_c(\rho, \mathcal{Q})$  if and only if the map  $\mathcal{Q}$  is invertible on all the pure states in the support of  $\rho$  [14], which for  $\rho = \mathbb{1}/d$  implies that the map is unitary.

### IV. CONCLUSIONS

We have introduced a new measure  $D(\rho, \{K_i\})$  of the disturbance (intended as an irreversible state change) that a map with Kraus operators  $\{K_i\}$  induces on a system in a state  $\rho$ . We have derived an information–disturbance tradeoff for such a quantity in the form  $I \leq D$ , where  $I$  is the classical info the map  $\{K_i\}$  returns on the state  $\rho$  to the experimenter. The equality conditions for this bound have been also derived. Moreover, a state-independent tradeoff  $\tilde{I} \leq \tilde{D}$  was obtained, which bounds the percentage of the maximum achievable information  $\tilde{I}$  with the disturbance  $\tilde{D}$  caused to a completely unknown input state.

Even though we found a valid information–disturbance tradeoff, we have to conclude that the disturbance definition  $D$  we used here may not be the appropriate definition to capture the true spirit of Heisenberg’s intuition (the “uncertainty principle”). In fact, a “classical” measurement (or a purely informational state change) where one gains information on which, out of a set of orthogonal configurations, our state is in, will perturb the state in an irreversible manner even though there is no dynamical interaction on the system. Heisenberg, on the other hand, analyzed only the irreversible state changes induced by dynamical actions (measurements) on the system. In this sense, the information–disturbance relation derived here might be considered as an extension of Heisenberg’s intuition on the disturbance a measurement induces.

### APPENDIX A: CONTINUITY OF $D$

Here we prove that the entropic disturbance  $D$  is continuous. More rigorously, we prove the following two statements: *i)*  $\rho \rightarrow \rho'$ , i.e.  $T(\rho, \rho') \rightarrow 0$ , implies  $D(\rho, \mathcal{Q}) \rightarrow D(\rho', \mathcal{Q})$ , where the trace distance  $T$  is defined as  $T(\rho, \rho') \equiv \text{Tr}[|\rho - \rho'|]/2$ ; *ii)*  $\mathcal{Q} \rightarrow \mathcal{Q}'$ , i.e.  $T(\mathcal{Q}[\rho], \mathcal{Q}'[\rho]) \rightarrow 0$ , implies  $D(\rho, \mathcal{Q}) \rightarrow D(\rho, \mathcal{Q}')$ .

Proof of *i*): Start from

$$D(\rho, \mathcal{Q}) - D(\rho', \mathcal{Q}) = [S(\rho) - S(\rho')] - [S(\mathcal{Q}(\rho)) - S(\mathcal{Q}(\rho'))] + [S_e(\rho, \mathcal{Q}) - S_e(\rho', \mathcal{Q})]. \quad (\text{A1})$$

The first bracket in Eq. (A1) tends to zero for  $\rho \rightarrow \rho'$  thanks to the continuity of the entropy. It derives from Fannes' inequality [20], according to which  $|S(\rho) - S(\rho')| \leq h(T(\rho, \rho'))$  with the function  $h(x) \rightarrow 0$  for  $x \rightarrow 0$ . The second bracket in Eq. (A1) analogously goes to zero since it is bounded by the first: The contractivity of CP-maps [21] implies that  $T(\mathcal{Q}(\rho), \mathcal{Q}(\rho')) \leq T(\rho, \rho')$ . To show that also the last bracket tends to zero, recall that the exchange entropy can be written as  $S_e(\rho, \mathcal{Q}) = S(W)$  with the matrix  $W$  defined by  $W_{ij} = \text{Tr}[K_i \rho K_j^\dagger]$ ,  $K_i$  being the Kraus operators of  $\mathcal{Q}$  [13]. Thus, Fannes' inequality

implies  $|S_e(\rho, \mathcal{Q}) - S_e(\rho', \mathcal{Q})| \leq h(T(W, W'))$ , and for  $\rho \rightarrow \rho'$ , we have  $W \rightarrow W'$ . In fact, since  $W$  is Hermitian,

$$\begin{aligned} |W_{ij} - W'_{ij}|^2 &= \text{Tr}[A_j^\dagger A_i (\rho - \rho')] \text{Tr}[A_i^\dagger A_j (\rho - \rho')] \\ &\leq \text{Tr}[A_j^\dagger A_i A_i^\dagger A_j] \text{Tr}[(\rho - \rho')^2], \end{aligned} \quad (\text{A2})$$

where we used the Schwarz inequality for the Hilbert-Schmidt scalar product of operators:  $\langle A|B \rangle \equiv \text{Tr}[A^\dagger B]$ .

Proof of *ii*): it follows immediately from the continuity of the entropy, i.e. from the Fannes inequality [20].

I thank G. Chiribella, G. M. D'Ariano and V. Giovannetti for stimulating hints and discussions. Financial support comes from MIUR through FIRB (bando 2001) and PRIN 2005.

- 
- [1] W. Heisenberg, *Zeitsch. Phys.* **43**, 172 (1927), English translation: J. A. Wheeler and H. Zurek, *Quantum Theory and Measurement* (Princeton Univ. Press, 1983), pg. 62.
  - [2] H. P. Robertson, *Phys. Rev.* **34**, 163 (1929).
  - [3] A. Peres *Quantum Theory: Concepts and Methods* (Kluwer ac. publ., Dordrecht, 1993).
  - [4] G. M. D'Ariano, *Fortschr. Phys.* **51**, 318 (2003).
  - [5] C. A. Fuchs and A. Peres, *Phys. Rev. A* **53**, 2038 (1996); C. A. Fuchs, *Fortschr. Phys.* **46**, 535 (1998); H. Barnum, *quant-ph/0205155* (2002); L. Maccone, *Phys. Rev. A* **73**, 042307 (2006).
  - [6] M. Ozawa, *Ann. Phys.* **311**, 350 (2004).
  - [7] C. H. Bennett, G. Brassard, and N. D. Mermin, *Phys. Rev. Lett.* **68**, 557 (1992); M. Ban, *J. Phys. A: Math. Gen.* **32**, 6527 (1999); K. Banaszek, *Phys. Rev. Lett.* **86**, 1366 (2001); P. Arrighi, *Int. J. Quantum Inf.* **2**, 341 (2004); F. Sciarrino, M. Ricci, F. De Martini, R. Filip, and L. Mišta, Jr. *Phys. Rev. Lett.* **96**, 020408 (2006); M. Christandl and A. Winter, *quant-ph/0501090* (2005).
  - [8] S. D. Bartlett, T. Rudolph, R. W. Spekkens, P. S. Turner, *New J. Phys.* **8**, 58 (2006).
  - [9] K. Kraus *States Effects and Operations* (Springer-Verlag, Berlin, 1983); K. Kraus, *Phys. Rev. D* **35**, 3070 (1987).
  - [10] R. Jozsa, D. Robb, and W. K. Wootters, *Phys. Rev. A* **49**, 668 (1994).
  - [11] F. Buscemi, G. M. D'Ariano, M. Keyl, P. Perinotti, and R. F. Werner, *J. of Math. Phys.* **46**, 082109 (2005).
  - [12] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2000).
  - [13] B. Schumacher, *Phys. Rev. A* **54**, 2614 (1996).
  - [14] B. Schumacher and M. A. Nielsen, *Phys. Rev. A* **54**, 2629 (1996).
  - [15] S. Lloyd, *Phys. Rev. A* **55**, 1613 (1997).
  - [16] A. S. Holevo, *IEEE Trans. Inf. Theory* **44**, 269 (1998); B. Schumacher and M. D. Westmoreland, *Phys. Rev. A* **56**, 131 (1997).
  - [17] H. P. Yuen, M. Ozawa, *Phys. Rev. Lett.* **70**, 363 (1992).
  - [18] H. Barnum, M. A. Nielsen, and B. Schumacher, *Phys. Rev. A* **57**, 4153 (1998); I. Devetak, *IEEE Trans. Inform. Theory* **51**, 44 (2005).
  - [19] M. J. W. Hall, *Phys. Rev. A* **55**, 100 (1997).
  - [20] M. Fannes, *Commun. Math. Phys.* **31**, 291 (1973).
  - [21] M. B. Ruskai, *Rev. Math. Phys.* **6**, 1147 (1994).
  - [22] Note that by "invertible" we mean that there exists a transformation that, acting on the final state of the system only, is able to recover the initial state even if the system was initially entangled with some other system. This means that if  $|\Psi\rangle$  is a purification of  $\rho$ , there exists an "inversion" CP-map identified by the Kraus operators  $\{I_j\}$  such that  $|\Psi\rangle\langle\Psi| = \sum_{ij} (\mathbb{1} \otimes I_j K_i) |\Psi\rangle\langle\Psi| (\mathbb{1} \otimes K_i^\dagger I_j^\dagger)$ , where the identity  $\mathbb{1}$  acts on the purification space.
  - [23] One could try enforcing requirements *ii*) and *iii*) by maximizing the fidelity over all possible unitary operators, defining a disturbance of the form  $\bar{D} = 1 - \max_{\bar{U}} F(\rho, \bar{U} \rho' \bar{U}^\dagger)$ , where  $F$  is the fidelity,  $\rho$  and  $\rho'$  are the input and output states, and the maximization runs over all unitaries  $\bar{U}$ . Also this definition is inadequate, since  $\bar{D}$  is null if  $\rho$  and  $\rho'$  have the same eigenvalues, which does not necessarily entail that the transformation is invertible, i.e. requirement *ii*) still does not hold.